

Methodische Beiträge

Datenqualität und Generalisierbarkeit von Online Access Panels: Reflexionen zur Datenbasis des Buchs

Dimitri Prandner¹ und Alexander Seymer^{1,2}

¹ JKU – Johannes Kepler Universität Linz

² PLUS - Paris Lodron Universität Salzburg

Kontakt : Dimitri.Prandner@jku.at

Abstract (dt.)

In diesem Sammelband sind 15 Beiträge von mehr als 20 Autor*innen enthalten, die sich mit den Herausforderungen und Folgen der Covid-19-Krise zwischen 2020 und 2022 beschäftigen. Die Beiträge bauen auf der *Values-in-Crisis*-Studie (VIC) mit Bevölkerungsumfragen auf, die mittels eines ISO-zertifizierten Online-Access-Panels und selbst auszufüllenden *Computer Assisted Web Interviews* (CAWI) realisiert wurde. Obwohl die Rekrutierung und Befragung von Teilnehmenden über das Internet mittlerweile gängige Praxis in Marktforschung und Wissenschaft sind, werden im Besonderen Online-Access-Panels als weniger zuverlässig für allgemeine Bevölkerungsumfragen betrachtet. Dies macht Detailanalysen über die Datenqualität notwendig. Tests der VIC-Daten zeigen aber, dass der Datensatz univariat in wesentlichen sozialstrukturellen Merkmalen nur geringe Abweichungen zur Grundgesamtheit hat und den Quotenvorgaben entspricht. Auch die *Anpassungsgewichtung* sowie der Ausstieg von Personen aus dem Panel, sogenannte *Panel Attrition*, liegen nur in vertretbaren Ausmaßen vor. Bezüglich des Antwortverhaltens ist durchgängig eine Tendenz zum schnellen Ausfüllen des Fragebogens beobachtbar, insbesondere bei komplexen Antwortbatterien wie der *PVQ-Skala*. Andere methodisch erfassbare Effekte wie Ja-Sage-Tendenzen sind inhaltlich plausibel und begründbar. Die Datengrundlage des vorliegenden Buches kann aber insgesamt als gut eingeschätzt werden, wobei die methodische Eignung für bestimmte Analysen wie bei jeder Studie im Einzelfall nachgeprüft und dargelegt werden muss.

Schlüsselwörter: Datenqualität, Datengewichtung, Antworttendenzen, Online Access-Panels

Abstract (eng.)

This anthology presents 15 chapters by more than 20 authors addressing the challenges and consequences of the Covid 19 crisis between 2020 and 2022. The contributions build on the *Values in Crisis* (VIC) population surveys, which were implemented using an ISO-certified online access panel and self-completion *Computer Assisted Web Interviews* (CAWI). Although recruiting and interviewing participants via the internet has become common practice in market research and academia, online access panels in particular are considered less reliable for general population surveys. This makes detailed analyses of data quality necessary. Tests of the VIC data show, however, that the data set has only minor deviations from the population in essential socio-structural characteristics and meets the quota requirements. The adjustment weighting as well as the exit of persons from the panel, so-called panel attrition, are also only present to an acceptable extent. With regard to response behaviour, a tendency to fill out the questionnaire quickly can be observed throughout, especially with complex answer batteries such as the PVQ scale.

Other methodological effects such as yes-saying tendencies are plausible and justifiable in terms of content. However, the data basis of this book can be assessed as good overall, although the methodological suitability for certain analyses must be verified and explained in each individual case, as is the case with every study.

Keywords: data quality, data weighting, response trends, online access panels

1 Einleitung

In diesem Sammelband finden sich 15 Beiträge von mehr als 20 Autor*innen, die sich auf Grundlage von quantitativen Umfragedaten mit den Einstellungen und Werten der österreichischen Wohnbevölkerung während der Covid-19-Pandemie zwischen 2020 bis 2022 beschäftigen. Was im Frühjahr 2020 ursprünglich als Gesundheitskrise begann, wurde im Verlauf der Jahre zu einer gesamtgesellschaftlichen Krise, die viele unterschiedliche Facetten hatte. Kam es kurz nach Beginn der Pandemie bereits zu Diskussionen über wirtschaftliche Aspekte und Debatten über die geeigneten Maßnahmen zur Pandemiebekämpfung, folgten weitere nationale und internationale Ereignisse, die die gesellschaftliche Lage weiter destabilisierten. Obwohl 2021 noch stark im Zeichen der Pandemie stand, kam es in Österreich beispielsweise zu den politischen Skandalen um die Volkspartei, die das Vertrauen in die Politik weiter erschüttert haben. 2022 stand im Schatten des Überfalls Russlands auf die Ukraine im Februar und der folgenden Energie- und Inflationskrisen, die neue Fragen aufwarfen.

Um solch vielfältige sozialwissenschaftliche Erkenntnisse über die Auswirkungen von Krisen auf das Leben der Österreicher*innen bereitzustellen, ist eine umfangreiche und belastbare empirische Datenbasis erforderlich (siehe Kittel et al., 2020; Prandner, 2022). In diesem Fall wurde der Datensatz der *Values-in-Crisis*-Studie (VIC; Aschauer et al., 2023) für die Analysen herangezogen. Jeweils im Frühsommer 2020, 2021 und 2022 wurden Daten über die Verfasstheit der österreichischen Bevölkerung gesammelt.

Die Datenerhebung während der Pandemiejahre war dabei eine besondere Herausforderung. Obwohl Krisen zu einem dringlichen Bedarf wissenschaftlicher Analysen führen, bringen die veränderten Rahmenbedingungen Probleme für die Erfassung von entsprechenden Informationen. Die Covid-19-Pandemie brachte im Wesentlichen zwei zentrale Herausforderungen für die Umfrageforschung mit sich: die schnelle Umsetzung von inhaltlich angepassten Befragungen zum einen und die staatlich verordneten Kontaktbeschränkungen zum anderen, die etablierte Datenerhebungsmethoden wie etwa persönliche Befragungen in Haushalten teilweise unmöglich machen.

Eine inhaltliche Anpassung von (teils bereits langfristig geplanten) Umfragen war notwendig, da die Pandemie ganz neue Fragestellungen in den Mittelpunkt rückte. Große Teile des öffentlichen Lebens wurden ausgesetzt. Gleichzeitig war aufgrund der schwer abschätzbaren Dauer der Pandemie

eine schnelle Datenerhebung notwendig, um die Krisen und ihre Folgen analysieren zu können. Repräsentative Bevölkerungsumfragen, welche zur kontinuierlichen Beobachtung von Dynamiken und Entwicklungen in der Gesellschaft angelegt sind, verwenden in der Regel aber kosten- und zeitintensive Verfahren zur Rekrutierung und Befragung von Personen, um eine hohe Datenqualität zu gewährleisten. Aufwändige Zufallsstichprobeneziehung oder Generierung von Auswahlgesamtheiten aus Registerdaten – genauso wie die häufig als Goldstandard betrachteten *Face-to-Face-Interviews* (F2F; vgl. Dillman 2020; Dillman et al., 2014) – nehmen in der Umsetzung viel Zeit in Anspruch und waren aufgrund der Kontaktbeschränkungen auch praktisch nicht realisierbar. Damit rückten während der Pandemie sowohl für die Rekrutierung als auch für die Erhebung Verfahren in den Fokus, welche eine zeitnahe Umsetzung ermöglichen, wobei hier besonders die Rekrutierung und Befragung über das Internet an Prominenz gewannen.

Auch die Studienteilnehmer*innen der österreichischen VIC-Studie wurden online rekrutiert. Dazu wurden auf Grundlage von Verteilungsquoten Teilnehmende aus einem Online-Access-Panel ausgewählt, die im Anschluss einen Fragebogen mittels Computer Assisted Web Interviews (CAWI) beantworteten. Die Rekrutierung und Befragung über das Internet sind keinesfalls neu, wurden aber lange Zeit als weniger zuverlässig für allgemeine Bevölkerungsumfragen betrachtet (Cornesse et al., 2020; Prandner et al., 2023). Deshalb soll in diesem Beitrag die Datenqualität entlang von drei zentralen Dimensionen diskutiert werden. Erstens wird die Rekrutierung der Befragten in die Stichprobe betrachtet, wobei die theoretischen Herausforderungen genauso wie die praktische Umsetzung über einen Vergleich von Rohdaten mit offiziellen Registerdaten der Statistik Austria betrachtet werden. Zweitens können die Stichproben untereinander verglichen werden, denn bei Paneldaten ist es wichtig zu eruieren, ob die Stichprobe auch über die unterschiedlichen Erhebungszeitpunkte von hoher Qualität bleibt. Drittens lassen sich anhand von spezifischen Analysen des Antwortverhaltens Schlüsse auf die Qualität der Daten ziehen. Das Fazit wird die drei Dimensionen in eine Gesamteinordnung der Datenqualität zusammenführen.

2 Die Stichprobe und der Vergleich mit Bevölkerungsstatistiken

Die Covid-19-Pandemie führte dazu, dass es zu einer Zunahme an Online-Umfragen kam, die als Ersatz zu etablierten F2F-Umfragen eingesetzt wurden. Besonders die Nutzung von sogenannten *Online-Access-Panels* stieg in der akademischen Umfrageforschung (Einarsson et al., 2022; Cornesse et al., 2022). Diese Angebote für die Online-Datenerhebung stützen sich auf freiwillige Nutzer*innen, die sich bei einem entsprechenden Panel registrieren und dann in regelmäßigen Abständen Einladungen zur Teilnahme an Studien erhalten (Blom et al., 2016; Keusch, 2015). Theoretisch haben diese Panels einen wesentlichen Vorteil gegenüber anderen Modi der Onlinedatensammlung: Es würde die Möglichkeit existieren, Wahrscheinlichkeitsstichproben aus dem Panel zu ziehen. Da jedoch nicht jede Person einer Grundgesamtheit – also z.B. alle Personen mit österreichischer Staatsbürgerschaft – die gleiche Chance hat, sich für ein solches Panel anzumelden (z. B. die für die Teilnahme erforderlichen digitalen Fähigkeiten, das Interesse an Studien oder das Ausmaß der Incentivierung für die Teilnahme), weisen die aus solchen Panels gezogenen Stichproben eine größere Auswahlverzerrung – in der Fachliteratur auch als Bias bezeichnet – auf als Zufallsstichproben, welche z.B. auf dem Wohnadressenregister oder dem zentralen Melderegister basieren.

Um dies zu korrigieren, werden aus Online-Access-Panels häufig sogenannte Quoten-Stichproben – also Nicht-Zufallsstichproben – gezogen, welche die sozialen Strukturmerkmale von Gesellschaften möglichst gut abbilden. Dies ist eine notwendige Voraussetzung für qualitativ hochwertige Umfrageforschung, da das Ziel der Analyse von quantitativen Daten grundsätzlich in einer Verallgemeinerung von Umfrageergebnissen liegt.

Diese sogenannte Strukturgleichheit (Erklärung siehe Abbildung 1 und Infobox 1) wird über die Quotierung ausgewählter Merkmale hergestellt. Beispielsweise wird eine vordefinierte Anzahl an Männern in einem bestimmten Alter rekrutiert. Eine schematische Darstellung und Erklärung für die Herausforderungen einer unverzerrten Abbildung der Grundgesamtheit in der Stichprobe kann in Abbildung 1 gefunden werden.

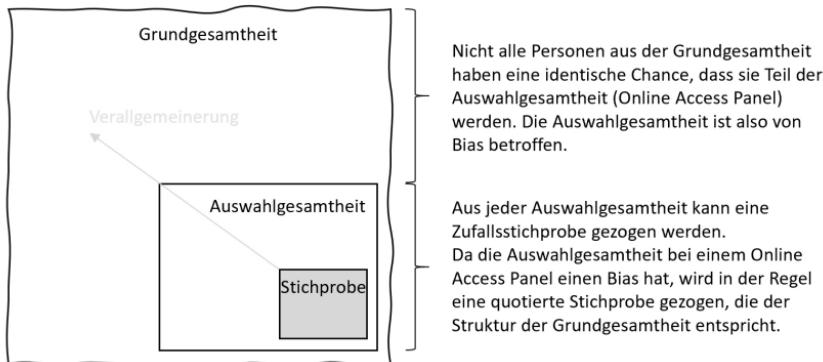


Abbildung 1: Von der Grundgesamtheit zur Stichprobe – die Auswahl der Befragten (eigene Abbildung)

Das Ziel der Umfrageforschung sind verallgemeinerbare Ergebnisse. Entsprechend muss es möglich sein, einen Bezug zwischen Grundgesamtheit und Stichprobe herzustellen. Voraussetzung dafür ist, dass die Grundgesamtheit durch die Stichprobe abgebildet werden kann, wobei zwischen strukturgleichen Stichproben und repräsentativen Stichproben zu unterscheiden ist.

- Strukturgleichheit bedeutet, dass sich die Verteilung von bekannten Merkmalen wie Altersstruktur, Geschlechterverteilung etc. zwischen Stichprobe und Grundgesamtheit deckt.
- Repräsentativität bedeutet, dass untersuchungsrelevante Informationen (z.B. das mittlere Einkommen) in einer Zufallsstichprobe nicht systematisch von den sogenannten wahren Werten der Grundgesamtheit abweichen.

Infobox 1: Strukturgleichheit und Repräsentativität

Sollten trotz Quotenverfahren die relevanten Strukturmerkmale aus der realisierten Stichprobe von der bekannten Verteilung innerhalb der untersuchten Gesellschaften abweichen, können die Daten gegebenenfalls gewichtet werden, um die Verteilung zu korrigieren (MacInnis et al., 2018; siehe auch der nächste Abschnitt). Diese Strategie ist in der Umfrageforschung üblich und stellt die Generalisierbarkeit der Angaben bzw. Aussagen der Befragten sicher, was in der Umfrageforschung ein typisches Qualitätskriterium für Daten darstellt (Ito & Todoroki, 2021).

Für Online-Access-Panels wurden auch professionelle Standards entwickelt, um Probleme bei der Rekrutierung, der Panelzusammensetzung und der Panelpflege zu lösen und die Qualität zu sichern. Es gibt eine entsprechende ISO-Norm für die Markt-, Meinungs- und Sozialforschung (International Organization for Standardization, 2019), an die sich seriöse Panelanbieter halten. Und mit *Marketagent.com online reSEARCH GmbH*

(kurz: Marketagent) wurde ein entsprechend zertifizierter Panelanbieter für die österreichische VIC-Studie beauftragt, der alle drei Erhebungswellen zwischen Mai 2020 und Juli 2022 durchführte.

Marketagent hat mehr als 2.300.000 eingeschriebene Panelteilnehmende in 75 Ländern, davon knapp 140.000 in Österreich (Marketagent, 2022). Neben der Zertifizierung sprach für die Wahl von Marketagent auch die Erfahrung des Anbieters mit sozialwissenschaftlichen Bevölkerungsstudien wie der österreichischen Wahlstudie *AUTNES* und das *Austrian Corona Panel Project* (siehe Kittel et al., 2020; Prandner, 2022).

Aus diesem Panel wurden 2.018 Personen im Mai 2020, 2.082 Personen zwischen März und April 2021 sowie 2.123 Personen im Juli 2022 rekrutiert, die auch den Fragebogen vollständig ausfüllten. Die Feldphase – also der Zeitraum der Datenerhebung, in dem Panelteilnehmende den Fragebogen ausfüllen konnten – erstreckte sich im Jahr 2020 über 10 Tage (beginnend mit dem 14. Mai), 2021 über 18 Tage (beginnend mit dem 25. März) und 2022 über 15 Tage (beginnend mit dem 5. Juli). Die Auswahl der Personen erfolgte auf Grundlage ihres Alters, Geschlechts, formalen Bildungsabschlusses und des Wohnortes (Bundesland) und das Ziel war, dass die Stichprobe strukturgleich zur österreichischen Wohnbevölkerung ab einem Alter von 14 Jahren sein sollte.

Bei der Prüfung der Voraussetzung der Strukturgleichheit der Stichprobe zeigt sich, dass die erhobenen Datensätze jeweils in weiten Teilen diese Anforderungen bereits vor Nutzung von sogenannten Anpassungsgewichtungen erfüllen konnten. Betrachtet man Alter und Geschlecht gemeinsam, ist ein Großteil der Abweichungen zwischen den Daten der Statistik Austria – die auf dem Melderegister beruhen – unauffällig und unter einem Prozentpunkt. Einzig Frauen über 70 Jahre wurden untererfasst. Während diese ca. 9% der Population ausmachen, war diese Gruppe in der Stichprobe jeweils nur mit ca. 6% vertreten. Diese Untererfassung war über alle drei Wellen hinweg konstant, was aber eine bekannte Konsequenz von Online-Datenerhebungen ist (Kittel et al., 2020).

Die einzigen relevanten Abweichungen in der Stichprobe sind im Bereich der formalen Bildungsabschlüsse zu finden, wenn man diese differenziert betrachtet. So sind Personen, die sehr niedrige bzw. hohe Bildungsabschlüsse vorweisen, durchwegs untererfasst worden, während Personen mit mittleren Bildungsabschlüssen überproportional häufig vorkommen.

Konkret bedeutet dies, dass die Personengruppe mit maximal einem Pflichtschulabschluss mit ca. 8 Prozentpunkten in Welle 1 und 11 Prozentpunkten in Welle 3 untererfasst wurde und Personen mit Hochschulab-

schluss zumindest in Welle 1 und 3 um ca. 4 Prozentpunkte untererfasst wurden. Dafür sind Personen, die einen Lehrabschluss vorweisen können, mit 7 Prozentpunkten in Welle 3 und knapp 11 Prozentpunkten in Welle 2 übererfasst worden bzw. jene, die eine BHS oder AHS besucht haben, mit ca. 3 Prozentpunkten in Welle 2 und 5 Prozentpunkten in Welle 3 übererfasst (siehe Tabelle 1).

Tabelle 1: Ungewichtete Rohdaten von Values-in-Crisis im Vergleich zur gemeldeten Bevölkerungsstatistik in Prozentpunkten

Eigenschaft	VIC - Welle 1	Statis-tik Austria 2020	Ab-wei-chung 2020	VIC - Welle 2	Statis-tik Austria 2021	Ab-wei-chung 2021	VIC - Welle 3	Statis-tik Austria 2022	Abwei-chung 2022
Frauen nach Altersgruppen	50,8	51,2	-0,4	50,5	51,1	-0,6	50,9	51	-0,1
	14-19 Jahre	3,1	3,3	-0,2	3,1	3,3	-0,2	3,3	3,3
	20-29 Jahre	8,1	7,1	1	8,2	6,9	1,3	8,3	6,8
	30-39 Jahre	7,9	7,8	0,1	8	7,8	0,2	8,1	7,8
	40-49 Jahre	8,7	7,7	1	8,4	7,6	0,8	8,6	7,6
	50-59 Jahre	9,4	9,1	0,3	9,3	9	0,3	9,6	9
	60-69 Jahre	7,2	6,8	0,4	7,3	7	0,3	7,4	7,2
	70+ Jahre	6,3	9,4	-3,1	6	9,4	-3,4	5,8	9,4
Männer nach Altersgruppen	49,2	48,8	0,4	49,5	48,9	0,6	49,1	49	0,1
	14-19 Jahre	2,9	3,5	-0,6	2,6	3,5	-0,9	2	3,5
	20-29 Jahre	8,5	7,5	1	8,4	7,3	1,1	8,5	7,2
	30-39 Jahre	7,8	8	-0,2	7,7	8	-0,3	8,1	8,1
	40-49 Jahre	8,4	7,7	0,7	8,4	7,6	0,8	8,5	7,6
	50-59 Jahre	9,9	9,1	0,8	9,6	9	0,6	9,5	8,9
	60-69 Jahre	6,2	6,3	-0,1	7,8	6,5	1,3	6,9	6,8
	70+ Jahre	5,5	6,8	-1,3	5,1	6,8	-1,7	5,6	6,9
Bundesland	Vorarlberg	4,4	4,5	-0,1	4	4,5	-0,5	3,9	4,4
	Tirol	8,6	8,5	0,1	8,1	8,5	-0,4	7,7	8,5
	Salzburg	6,2	6,3	-0,1	6,3	6,3	0	5	6,2
	Steiermark	14,3	14	0,3	14,3	14	0,3	14,5	14,1
	Kärnten	6,5	6,3	0,2	6,3	6,3	0	5,9	6,4
	Oberöster-reich	16,9	16,7	0,2	17,5	16,7	0,8	15,8	16,6
	Niederös-terreich	18,9	18,9	0	18,9	18,9	0	21,3	18,9

Eigenschaft	VIC - Welle 1	Statistik Austria 2020	Abweichung 2020	VIC - Welle 2	Statistik Austria 2021	Abweichung 2021	VIC - Welle 3	Statistik Austria 2022	Abweichung 2022	
Formaler Bildungabschluss	Wien	20,5	21,5	-1	20,9	21,5	-0,6	22,4	21,5	0,9
	Burgenland	3,7	3,3	0,4	3,6	3,3	0,3	3,4	3,4	0
	Pflichtschule	17,2	24,8	-7,6	14	24,8	-10,8	13,7	24,8	-11,1
	Lehre	38,7	31,1	7,6	41,8	31,1	10,7	38,1	31,1	7
	Berufsbildende mittlere Schulen	13,1	13,8	-0,7	14,8	13,8	1	13,9	13,8	0,1
	Allgemeine oder berufsbildende höhere Schule (AHS / BHS)	19,7	15,5	4,2	18,7	15,5	3,2	20,3	15,5	4,8
	Universität bzw. andere Hochschule	11,4	14,9	-3,5	10,8	14,9	-4,1	14	14,9	-0,9

Anmerkungen: eigene Darstellung basierend auf Tabelle 2 im Methodenbericht (Aschauer et al., 2023, S. 6)

3 Vergleich der Teilstichproben im Panel

Eine weitere Möglichkeit der Beurteilung der Qualität von Paneldaten wird durch den Vergleich von Teilstichproben entlang der Teilnahme im Panel möglich. In Wiederholungsbefragungen nehmen nicht alle Befragten an allen Wellen einer Studie teil. Im Fall der VIC-Studie hatte das die Konsequenz, dass Befragte entweder nur an einer Befragung teilgenommen haben (W1, W2, W3) oder an zwei Wellen (W12, W23, W13) oder an allen drei Wellen (W123). Idealerweise unterscheiden sich diese sieben Teilstichproben nicht in der Zusammensetzung und man könnte von einem zufälligen Ausfall von Panelteilnehmenden sprechen. Im Folgenden wird dies über die Gewichtung und über die Verteilung von soziodemographischen Merkmalen ausdifferenziert.

Die für die VIC-Studie berechneten Gewichte wurden über einen Raking Algorithmus (Pasek, 2022) erstellt, um die Verteilung der Umfragedaten an

die offiziellen Verteilungsangaben der Statistik Austria bezüglich der Merkmale Geschlecht, Alter, Bildung und Region anzupassen. Personengruppen, welche genauso oft im Datensatz vorkommen, wie es die Verteilung der offiziellen Statistik nahelegt, erhalten ein Gewicht von 1. Personen mit Merkmalen, die zu häufig in der Stichprobe vertreten sind, erhalten ein Gewicht von kleiner 1 und Personen, die unterrepräsentiert sind, erhalten ein Gewicht größer 1. Grundsätzlich ist die Gewichtung nur für Gewichte mit Werten größer als 1 tendenziell problematisch, weil das Antwortverhalten der Person stärker gewichtet wird, was einerseits mit einer Verringerung der Varianz – also der Streuung der Antworten um den jeweiligen Mittelwert – einhergeht und andererseits auch mögliche inhaltliche Verzerrungen mit sich bringt. Oftmals werden Gewichte ab einem Wert von 4 als unverhältnismäßig groß beschrieben (Gabler & Ganninger, 2010) und alle Gewichte, die über einen Faktor von 4 hinausgehen, werden gestrichen (Little et al., 1997).

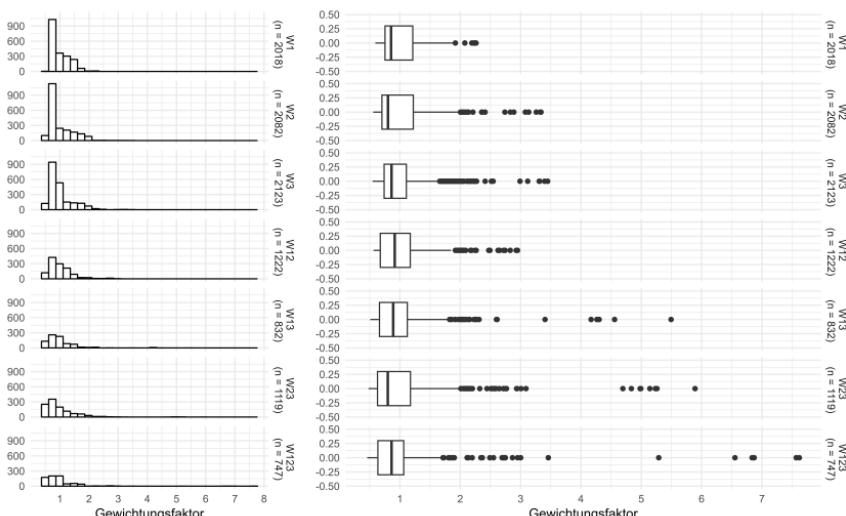


Abbildung 2: Histogramm und Boxplot für Gewichtungswerte nach Stichprobe (eigene Darstellung; Datenquelle: VIC 1-3)

Für die vorliegende Beurteilung der Daten ist die Verteilung der Gewichte auch jenseits von Grenzwerten interessant, um den Grad der Abweichung zu beurteilen. Aus Abbildung 2 wird deutlich, dass für die einzelnen Wellen 1 bis 3 (W1, W2, W3) keine relevanten Abweichungen erkennbar sind. Auch

für die Teilnehmer*innen der Panelbefragung, die an den Wellen 1 und 2 (Gruppe: W12) teilgenommen haben, ist keine Anpassung nötig. Erst die übrigen Paneldaten deuten auf eine geringe Anzahl an Fällen mit Gewichten jenseits von 4 hin. Die Paneldaten über alle drei Wellen zeigen dabei die höchste Abweichung von den Registerdaten. Aschauer et al. (2023) weisen nichtsdestotrotz im Methodenreport zur Studie maximal 12 Personen aus, die Gewichte mit einem Wert von größer als 3 haben. Vergleicht man dies mit den großen Stichprobenumfängen, ist diese Anzahl vernachlässigbar klein.

Nach dem Vergleich der Gesamtstichproben pro Welle mit den Registerdaten und der Betrachtung der Gewichte soll nun noch ein detaillierter Blick auf die sogenannte *Panel Attrition* erfolgen, getrennt nach den Gewichtungsmerkmalen Geschlecht, Alter, Bildung und Region. Als *Panel Attrition* wird das Ausscheiden von Befragten über die Wellen bezeichnet. Idealerweise würden alle Befragte an allen Wellen teilnehmen. Allerdings lässt sich nur ein begrenzter Personenkreis für eine erneute Befragung motivieren. Grundsätzlich ist Panel Attrition unproblematisch, solange das Ausscheiden der Personen zufällig ist. Diese Zufälligkeit würde sich dann darin widerspiegeln, dass sich die Verteilungen auf den soziodemographischen Variablen, wie Geschlecht, Alter, Bildung oder Region, für die unterschiedlichen Substichproben des Gesamtpans nicht unterscheiden. Deshalb wird die Panelstichprobe unterteilt in Substichproben von Personen. So kann klar identifiziert werden, ob Personen nur an einer Welle (W1, W2, W3), an zwei Wellen (W12, W13, W23) oder an allen drei Befragungswellen (W123) teilgenommen haben.

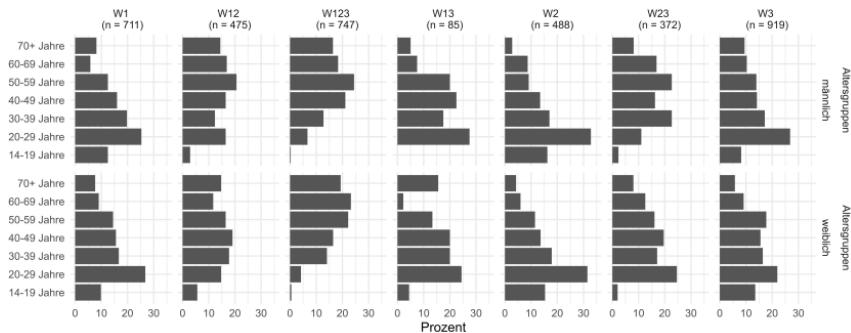


Abbildung 3: Häufigkeitsverteilung von Alter, getrennt nach Geschlecht, für die unterschiedlichen Substichproben (n = gültige Fälle; eigene Darstellung; Datenquelle: VIC 1-3).

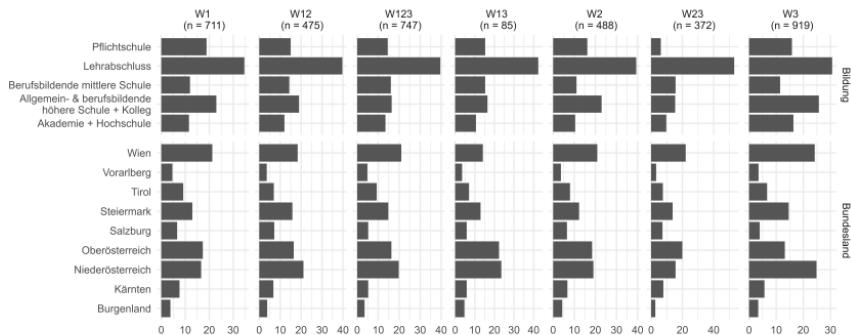


Abbildung 4: Häufigkeitsverteilung für Bildung und Region nach den unterschiedlichen Substichproben (n = gültige Fälle; eigene Darstellung; Datenquelle: VIC 1-3).

Abbildung 3 und Abbildung 4 zeigen die Häufigkeitsverteilungen für die unterschiedlichen Substichproben der Panelteilnehmenden, unterteilt in die soziodemographischen Merkmale, die zur Stichprobenziehung herangezogen wurden. Bei einem vollkommen zufälligen Ausscheiden aus dem Panel würden die Verteilungen in jeder Zeile komplett identisch aussehen, wobei die absoluten Häufigkeiten variieren können. Während sich in Abbildung 4 die Verteilung für Bildung und Bundesland verhältnismäßig einheitlich darstellt, sind in Abbildung 3 deutliche Unterschiede zwischen den Verteilungen in den jeweiligen Gruppen erkennbar. In der Gruppe von

Personen, die an allen drei Wellen teilgenommen haben (Gruppe: W123), sind die jüngeren Altersgruppen bis 40 Jahre deutlich unterrepräsentiert. Dies zeigt sich auch in deutlich höheren Häufigkeiten dieser Altersgruppen, wenn man die einzelnen Teilnehmer*innengruppen für die Wellen 1 bis 3 (W1, W2 und W3) betrachtet. Denn hier wurde die Stichprobe über die (Nach-)Rekrutierung angepasst und somit sind in diesen Teilstichproben, welche aus Personen mit einmaliger Teilnahme bestehen, die jüngeren Altersgruppen stärker vertreten. Die Dynamiken sind für Männer und Frauen ähnlich. Der Alterseffekt spiegelt sich auch in der Bildung wider, wo Personen mit Lehrabschluss im Panel überrepräsentiert sind. Beim Bundesland ist auffallend, dass Niederösterreich und Wien in der dritten Welle (W3) besonders stark rekrutiert wurden, was wiederum bedeutet, dass Personen aus Niederösterreich und Wien stärker von *Panel Attrition* betroffen sind.

4 Antwortverhalten – Analyse von Messfehlern

Nachdem in den letzten Abschnitten dargelegt werden konnte, dass sich die Stichprobe nur in ausgewählten Bereichen von der Grundgesamtheit unterscheidet und auch Abweichungen bei den Teilstichproben nur in einem vertretbaren Ausmaß vorkommen, wird im kommenden Abschnitt darauf eingegangen, welche Rückschlüsse auf die Datenqualität durch das Antwortverhalten der Befragten gezogen werden können.

Bei der Erhebung von subjektiven Einstellungen, Werten und Meinungen, denen keinen objektiven Werte zu Grunde liegen, ist die Gefahr groß, dass die im Rahmen von Umfragen erhobenen Daten durch individuelle Antworttendenzen und systematische Antwortmuster verzerrt werden können. Entsprechend ist es notwendig, Umfragedaten auf solche potenziellen Fehlerquellen zu überprüfen (Moosbrugger & Brandt, 2020; Johnson et al., 2011). In der Literatur wird in der Regel auf Messfehler eingegangen, die mit bestimmten Antwortstilen in Zusammenhang stehen und sich wie folgt klassifizieren lassen (Johnson et al., 2011):

- a. *Straightlining*: die Tendenz, unabhängig vom Frageinhalt wiederholt identische Antwortkategorien zu wählen.
- b. *Extremantworten*: die Tendenz, Randkategorien auf Antwortskalen zu wählen.
- c. *Akquieszenz*: die Tendenz, dass Aussagen – unabhängig von der inhaltlichen Ausrichtung – zugestimmt wird.

- d. *Speeding*: die schnelle, unreflektierte Beantwortung von Fragen, wobei die Fragestellung bzw. Antwortmöglichkeiten nicht (genau) gelesen werden.

Für die nun folgende Analyse wird auf mehrere sogenannte Fragebatterien – also inhaltlich ähnliche Aussagen, die auf Grundlage gleichlautender Antwortkategorien beantwortet werden – zurückgegriffen, die in den Wellen 2 und 3 der Values-in-Crisis-Studie zum Einsatz kamen. Auf Analysen zu Welle 1 wird verzichtet, da die entsprechenden Daten bereits von Prandner (2022, 342f) auf Antwortstile analysiert und keine substanzielles Einflüsse vorgefunden wurden.

Für die folgenden Analysen werden drei Fragebatterien herangezogen, für die im Rahmen der Erhebung auch Zeitmessungen stattgefunden haben. Diese liegen für einzelne Fragebatterien als sekundengenaue Paradata¹ vor und protokollierten die Antwortzeit der jeweiligen Befragten.

- a. Die erste erfasste Batterie ist die Schwartz *Portrait Value Questionnaire (PVQ)*-Skala, zur Ermittlung grundlegender Werteorientierungen von Shalom Schwartz (2012): Die Skala umfasst 21 Aussagen, die fiktive Personen vorstellen. Die Befragten beurteilen, von 1 – *sehr ähnlich* bis 6 – *gar nicht ähnlich, wie ähnlich sie diesen fiktiven Personen sind*. Die Skala wurde für die VIC-Erhebungen 2021 und 2022 analysiert und umfasst 1.021 zu lesende Worte. Eine Einführung in das Konzept der Wertemessung mittels PVQ kann in Prandner (2023) nachgelesen werden.
- b. Die zweite erfasste Fragebatterie umfasst 11 Aussagen, die unterschiedliche Haltungen zur Covid-19-Krise illustrieren. Die Batterie ist Teil der VIC-Studie aus dem Jahr 2022 und die Antworten reichen von 1 – *stimme gar nicht zu* bis 4 – *stimme vollkommen zu*. Die gesamte Skala inklusive Antworten umfasst 311 Worte. Die Batterie wurde selbst von der Studienleitung entwickelt und bereits in der Salzburger Autoritarismusstudie (Aschauer & Heinz, 2023) eingesetzt.
- c. Die dritte getestete Batterie umfasst das Konstrukt der gruppenbezogenen Menschenfeindlichkeit. Die Batterie umfasst 12 Aussagen mit Antworten von 1 – *stimme voll und ganz zu* bis 5 – *stimme überhaupt nicht zu*. Insgesamt hat die Skala 264 zu lesende Worte. Das Konzept und die

1 Paradata: Daten und Kontextinformationen zur Beschreibung der erhobenen Umfrageten. Beispiele dafür können der Ort der Befragung sein, die Befragungsdauer, die Zahl der anwesenden Personen während der Befragung etc. (siehe auch Prandner, 2019)

Fragestellungen gehen auf den deutschen Sozialwissenschaftler Wilhelm Heitmeyer (2005) zurück und das Konzept wurde von Aschauer (2020) rezent im österreichischen Kontext angewandt. Im vorliegenden Band wurden die Fragen zum Teil im Beitrag von Heinz verwendet.

Die Klassifizierung der Antwortstile *Akquieszenz*-Tendenz und *Extrem*-Antwort-Tendenz basiert auf den Ausführungen von Johnson und Kolleg*innen (2011). Entsprechend wird davon ausgegangen, dass bei mindestens 50% Übereinstimmung einer Batterie mit einem bestimmten Antwortverhalten ein Hinweis auf den jeweiligen Antwortstil vorliegt, – also z.B., wenn von 22 Fragen aus der Schwartz-Werteskala zumindest 11 Fragen mit einer Extremkategorie von *1 – sehr ähnlich* oder *6 – gar nicht ähnlich* beantwortet wurden. Wenn zumindest 75% der Antworten einem Stil entsprechen, liegt eine Tendenz zu dem Antwortstil vor – also z.B., wenn zumindest bei neun von elf Aussagen zur Covid-19-Krise eine identische Beurteilung erfolgte. Wenn alle Indikatoren einer Skala gleich beantwortet werden (und somit eine 100% Übereinstimmung gegeben ist), existiert der klare Nachweis für einen Antwortstil. Diese Klassifizierung entspricht der Vorgehensweise der Analysen für die erste Welle der VIC-Studie (Prandner, 2022).

Für das Konzept des *Straightlining* wurde die Varianz der Antworten auf die jeweiligen Aussagen herangezogen: Zeigen die gegebenen Antworten einer befragten Person keine Varianz – also keine Unterschiede – liegt *Straightlining* vor. Dies würde bedeuten, dass beispielsweise alle Aussagen bei der PVQ-Skala mit einer *2 – eher ähnlich* beantwortet sind.

Zur Messung des *Speedings* wurde auf Grundlage der Metaanalyse von Brysbaert (2019) die durchschnittliche zu erwartende Lesedauer für die entsprechenden Fragebatterien berechnet. In der Übersicht von Brysbaert wurde die durchschnittliche, nicht verbalisierte Lesegeschwindigkeit in Deutsch mit 260 Worten pro Minute festgehalten (Brysbaert, 2019, 10). Um die Komplexität der Fragebeantwortung auszugleichen, wurden 20 Wörter pro Minute abgezogen und die durchschnittliche Lesegeschwindigkeit mit 240 Worten pro Minute festgelegt. Personen werden als *Speeder* klassifiziert, wenn sie weniger als 40% der durchschnittlichen Lesezeit für das Ausfüllen einer Batterie benötigten. Personen gelten als schnell Ausfüllende, wenn sie zwischen 41% und 80% der durchschnittlichen Lesezeit liegen. Und jene, die sich zwischen 81% und 120% der durchschnittlichen Lesezeit befinden, werden als normal Ausfüllende eingeschätzt. Langsam Ausfüllende sind mit einer durchschnittlichen Lesezeit zwischen 121% und 160%

definiert und Personen, welche mehr als 160% der durchschnittlichen Lesezeit benötigen, gelten als sehr langsam Ausfüllende.

Betrachtet man die Ergebnisse dieser Analysen für die PVQ-Skala von Schwartz, zeigt sich ein konsistentes Bild. Da die PVQ-Skala mit 21 Aussagen vergleichsweise umfangreich ausfällt und Bestandteil aller drei Wellen der Values-in-Crisis-Studie ist, verwundert es wenig, dass eine größere Anzahl von *Speedern* beobachtet werden kann. Sind es in der Erhebung von 2021 ca. 25%, können 2022 sogar mehr als 30% der Befragten als *Speeder* klassifiziert werden. Zusätzlich zeigt sich, dass 2021 zwei Drittel und 2022 ca. 71% der Befragten bei diesen Frageblöcken zumindest als schnell Ausfüllende klassifiziert werden. Aufgrund der ähnlichen Struktur aller Aussagen in der PVQ-Skala und des Umfangs des Frageblocks kann dies aber auch als Konsequenz eines Lerneffekts gesehen werden. Die Befragten sind zu dem Zeitpunkt, als die Batterie im Fragebogen vorkam, bereits mit der Art der Fragestellung und der Struktur der Antwortmöglichkeiten vertraut gewesen. Da weder beim *Straightlining* (1,1% in 2021 und 1,8% in 2022) noch bei den Extremantworten oder der Akquieszenz besondere Auffälligkeiten bei diesen Fragen zu finden waren, sollte hier in zukünftigen Untersuchungen besonderer Wert darauf gelegt werden z.B. Kontrollfragen für Plausibilitätstests einzubauen. Es ist aber davon auszugehen, dass die Antworten inhaltlich konsistent sind.

Betrachtet man die Aussagen zu den Covid-19-Verschwörungserzählungen, zeigt sich ein etwas anderes Bild. In dieser Batterie sind 12% der Beantwortenden als *Speeder* zu klassifizieren. Aber mit knapp 6,5% Personen, die *Straightlining* betrieben, und mit mehr als 40%, die Anzeichen für Extremantwortverhalten zeigen, müssen diese Ergebnisse durchaus detaillierter betrachtet werden. So ist hier der Inhalt der Fragen relevant, da Aussagen wie „Die verfügbaren Impfstoffe gegen das Corona-Virus schaden mehr als sie nutzen“ durchaus polarisierend sind und z.B. Wissenschaftsskepsis oder den Glauben an Verschwörungserzählungen testen. Entsprechend sind kontinuierlich starke Zustimmungen oder Ablehnungen plausibel. Die Effekte, die man hier auf methodischer Ebene sieht, sind also inhaltlich durchaus nachvollziehbar.

Bei den Fragen zur Messung der gruppenspezifischen Menschenfeindlichkeit liegt wiederum ein differenziertes Bild vor. Einerseits sind hier bei mehr als 20% der Befragten Tendenzen zu Extremantworten zu finden – andererseits zeigt sich hier ein hoher Anteil an Personen (42%), die die Fragen zumindest langsam, wenn nicht sogar sehr langsam beantwortet haben. Entsprechend ist davon auszugehen, dass die Befragten über ihre

Tabelle 2: Antwortverhalten bei ausgewählten Fragebatterien

Antwortverhalten	Klassifizierung	Schwartz PVQ – 21 (N = 2082)	Schwartz PVQ – 22 (N = 2123)	Covid-19-Verschwörungserzählungen -22 (N = 2123)	Gruppen-spezifische Menschenfeindlichkeit -22 (N = 2123)
Speeding	Speeder	25,6%	30,6%	12,2%	14,8%
	schnell Ausfüllende	41,7%	40,9%	31,7%	10,8%
	normal Ausfüllende	20,9%	17,9%	32,7%	31,1%
	langsam Ausfüllende	5,7%	4,3%	12,6%	22,6%
Straightlining	sehr langsam Ausfüllende	6,1%	6,3%	10,8%	20,8%
	Vollkommenes Straightlining (keine Varianz in den Antworten)	1,1%	1,8%	6,5%	2,0%
Extremant-worten	Anzeichen (50% +)	7,6%	8,1%	43,3%	21,5%
	Tendenz (75% +)	1,6%	2,1%	22,0%	6,9%
	vollkommen (100%)	0,5%	1,0%	5,5%	1,9%
Akquieszenz	Anzeichen (50% +)	3,9%	3,5%	6,0%	1,4%
	Tendenz (75% +)	0,8%	0,7%	0,9%	0,2%
	vollkommen (100%)	0,1%	0,4%	0,6%	0,1%

Anmerkung: Datenquelle: VIC 2 und 3, eigene Darstellung

Antworten zu komplexen Fragen, wie der Bedeutung von Zuwanderenden für die nationale Identität (z.B. mit der Aussage: „Die in Österreich lebenden Zuwanderer sind eine Bereicherung für die österreichische Kultur“), nachgedacht und reflektiert haben.

Insgesamt gibt es neben der Tendenz, dass die lange Fragebatterie der PVQ-Skala von vielen der Befragten sehr schnell beantwortet wurde, keine außergewöhnlichen Auffälligkeiten im Datensatz (siehe Tabelle 2).

5 Fazit zur Datenqualität der VIC-Studie

Mit dem Beginn der Covid-19-Pandemie im Jahr 2020 sah sich die Umfrageforschung vor große Herausforderungen gestellt. Der Bedarf, innerhalb kurzer Zeit wissenschaftliche Daten zu generieren, stand ursprünglich den situationsbedingten Einschränkungen bei der Datenerhebung gegenüber. Im Rahmen pragmatischer Überlegungen akzeptierten Forschende bei einer Vielzahl an Studien diese Einschränkungen, wie eben bei Online-Ac-

cess-Panels, die bei der Value-in-Crisis-Studie und nicht nur dort eingesetzt wurden (siehe z.B. die *EVS-Sondererhebung* aus 2021-22 bei Kitzinger et al., 2023) oder die *Digital Skills Austria 2022* Studie (Grünangerl & Prandner, 2022).

Die Analysen in dem Kapitel zeigen, dass durch das Vorgehen über alle drei Wellen der VIC-Studie weitestgehend Strukturgleichheit mit der österreichischen Bevölkerung erreicht werden konnte. Dennoch gibt es in den Befunden Indizien, dass vereinzelt Abweichungen auftreten: So sind einerseits insbesondere ältere Frauen in der Stichprobe unterrepräsentiert und junge Personen konnten seltener motiviert werden, im Panel zu bleiben. Diese Gruppe musste häufiger nachrekrutiert werden. Bei den Antworttendenzen zeigt sich zusätzlich, dass es insbesondere bei der PVQ-Skala zu einem sehr schnellen Ausfüllen durch die Befragten kam.

Insgesamt kann in diesem Kapitel, das Stichproben- und auch Messqualität untersucht, aufgezeigt werden, dass die Einschränkungen von Online-Access-Panel-Daten, die in Kombination mit Quotenverfahren generiert wurden, reflektiert werden müssen. Denn die zunehmende Verfügbarkeit von Internetzugängen und steigender Kompetenz im Umgang mit Online-Medien und -Plattformen hat zwar dazu geführt, dass weite Bevölkerungsschichten kostengünstig und schnell für Umfragen erreicht werden können (Vehovar & Lozar Manfreda, 2017); aber Fragen über Rekrutierung, Antwortqualität und Generalisierbarkeit sind weiterhin zentral für die Einschätzung der Umfragequalität. Dabei darf nicht vergessen werden, dass die Online-Bevölkerung – bzw. der Teil von ihr, der im Online-Access-Panel von *Marketagent* registriert ist – nicht notwendigerweise der Wohnbevölkerung entspricht. Auch wenn die Strukturgleichheit bei VIC auf univariater Ebene in den meisten Fällen erreicht werden konnte, muss auf Limitierungen hinsichtlich der Verallgemeinerbarkeit von inhaltlichen Ergebnissen hingewiesen werden, wie Carina Cornesse und ihre Kolleg*innen (2020) auch in ihrer Abhandlung über unterschiedliche Strategien der Teilnehmendenrekrutierung dargelegt haben.

Dementsprechend sollte die vorliegende Abhandlung über die Methodik der *Values-in-Crisis*-Studie in Österreich als Beispiel gesehen werden, wie sich die Generierung von empirischen Daten an besondere Umstände anpasst. Die hier präsentierten Analysen sind nur ein Teil der möglichen Qualitätskontrollen von Umfragedaten; und durch detaillierte Erfassung von Daten zu Befragungsabbruch, genutzten Geräten zum Bearbeiten von Umfragen und anderen Paradaten (z.B. die Angaben, die bei der Registrierung für das Panel angegeben wurden) können in zukünftigen oder wei-

terführenden Studien noch genauere methodische Analysen durchgeführt werden. Denn mit reflektierten methodologischen Betrachtungen und mit einem klaren Hinweis auf Limitationen sind inhaltlich relevante und generalisierbare Befunde auf der Basis von Online-Access-Panel-Daten möglich.

Literatur

- Aschauer, W. (2020). *Religiöse Einstellungen als Einflussfaktor für Vorurteile oder Toleranz? Eine Analyse islamkritischer Haltungen in Österreich*. Österreichische Zeitschrift für Soziologie, 45, 183-212.
- Aschauer, W. & Heinz, J. (2023). *Autoritäre Dynamiken in der Corona-Krise. Einblicke in das Ausmaß von Autoritarismus und Coronaskepsis in der Salzburger Bevölkerung*. JBZ-Arbeitspapier 64, 2023.
- Aschauer, W., Seymer, A., Ulrich, M., Kreuzberger, M., Bodi-Fernandez, O., Höllinger, F., Eder, A., Prandner, D., Hadler, M. & Bacher, J. (2023). *Values in Crisis Austria – Wave 1, Wave 2 and Wave 3 combined (SUF edition)*. AUSSDA. <https://doi.org/10.11587/EYJMEZ> (Stand: 15.1.2024).
- Blom, A. G., Bosnjak, M., Cornilleau, A., Cousteaux, A. S., Das, M., Douhou, S. & Krieger, U. (2016). *A comparison of four probability-based online and mixed-mode panels in Europe*. Social Science Computer Review, 34(1), 8-25.
- Brysbaert, M. (2019). *How many words do we read per minute? A review and meta-analysis of reading rate*. Journal of Memory and Language, 109, 104047.
- Cornes, C., Blom, A. G., Dutwin, D., Krosnick, J. A., De Leeuw, E. D., Legleye, S., Pasek, J., Pennay, D., Phillips, B., Sakshaug, J. W., Struminskaya, B. & Wenz, A. (2020). *A review of conceptual approaches and empirical evidence on probability and nonprobability sample survey research*. Journal of Survey Statistics and Methodology, 8(1), 4-36. <https://doi.org/10.1093/jssam/smy028> (Stand: 15.1.2024).
- Cornes, C., Krieger, U., Sohn, M. L., Fikel, M., Friedel, S., Rettig, T., ... & Blom, A. G. (2022). *From German Internet Panel to Mannheim Corona Study: Adaptable probability-based online panel infrastructures during the pandemic*. Journal of the Royal Statistical Society Series A: Statistics in Society, 185(3), 773-797. <https://doi.org/10.1111/rssa.12644> (Stand: 15.1.2024).
- Dillman, D. A. (2020). *Towards survey response rate theories that no longer pass each other like strangers in the night*. In Understanding survey methodology: Sociological theory and applications. Routledge, 15-44.
- Dillman, D. A., Smyth, J. D. & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. John Wiley und Sons.
- Einarsson, H., Sakshaug, J. W., Cernat, A., Cornes, C. & Blom, A. G. (2022). *Measurement equivalence in probability and nonprobability online panels*. International Journal of Market Research, 64(4), 484-505. <https://doi.org/10.1177/1470785321105314> (Stand: 15.1.2024).

- Gabler, S. & Ganninger, M. (2010). Gewichtung. In C. Wolf und H. Best (Eds.), Handbuch der sozialwissenschaftlichen Datenanalyse. VS Verlag für Sozialwissenschaften, 143-164.
- Grünangerl, M. & Prandner, D. (2022). *Digital Skills Austria 2022*. RTR - Rundfunk und Telekom Regulierungs-GmbH. <https://doi.org/10.5281/zenodo.7333304> (Stand: 15.1.2024).
- Heitmeyer, W. (2005). *Gruppenbezogene Menschenfeindlichkeit. Die theoretische Konzeption und empirische Ergebnisse aus 2002, 2003 und 2004*. In W. Heitmeyer (Hrsg.), Deutsche Zustände, Folge, 3. VS Verlag für Sozialwissenschaften, 13-34.
- Ito, D. & Todoroki, M. (2021). *Evaluating the quality of online survey data collected in 2018 in the USA: Univariate, bivariate, and multivariate analyses*. International Journal of Japanese Sociology, 30(1), 140-162. <https://doi.org/10.1080/0911125X.2021.1890647> (Stand: 15.1.2024).
- International Organization for Standardization (2019). *ISO 20252:2019 Market, opinion and social research – Vocabulary and service requirements*. <https://www.iso.org/standard/72207.html> (Stand: 15.1.2024).
- Johnson, T. P., Shavitt, S. & Holbrook, A. L. (2011). *Survey response styles across cultures*. In D. Matsumoto und F. J. R. van de Vijver (Eds.), Culture and psychology. Cross-cultural research methods in psychology. Cambridge University Press, 130-175.
- Keusch, F. (2015). *Why do people participate in Web surveys? Applying survey participation theory to Internet survey data collection*. Management review quarterly, 65(3), 183-216.
- Kittel, B., Kritzinger, S., Boomgaarden, H., Prainsack, B., Eberl, J. M., Kalleitner, F., Lebernegg, N., Parthemüller, J., Plescia, C., Schiestl, D. & Schlogl, L. (2020a). *The Austrian Corona Panel Project: Monitoring individual and societal dynamics amidst the COVID-19 crisis*. European Political Science, 1-27. <https://doi.org/10.1057/s41304-020-00294-7> (Stand: 15.1.2024).
- Kritzinger, S., Willmann, J., Rohs, P., Pollak, M., Friesl, C. & Polak, R. (2023). *The European Values Study - Austrian COVID-19 Special Edition 2021-22 Including Youth Oversample (SUF edition)*. AUSSDA. <https://doi.org/10.11587/F8V4GL> (Stand: 15.1.2024).
- Little, R. J. A., Lewitzky, S., Heeringa, S., Lepkowski, J. & Kessler, R. C. (1997). *Assessment of Weighting Methodology for the National Comorbidity Survey*. American Journal of Epidemiology, 146(5), 439–449. <https://doi.org/10.1093/oxfordjournals.aje.a009297> (Stand: 15.1.2024).
- MacInnis, B., Krosnick, J. A., Ho, A. S. & Cho, M. J. (2018). *The accuracy of measurements with probability and nonprobability survey samples: Replication and extension*. Public Opinion Quarterly, 82(4), 707-744.
- Marketagent (2022). *Panelbook 2022*. <https://b2b.marketagent.com/media/5jrj43nr/marketagent-panel-book-2022.pdf> (Stand: 15.1.2024).
- Moosbrugger, H. & Brandt, H. (2020). *Antwortformate und Itemtypen*. In H. Moosbrugger und A. Kelava (Eds.), Testtheorie und Fragebogenkonstruktion. Springer, 91-117.

- Pasek, J. (2022). *anesrake: ANES Raking Implementation*. <https://cran.r-project.org/web/packages/anesrake/anesrake.pdf> (Stand: 15.1.2024).
- Prandner, D. (2019). *Zur Methodik des Sozialen Survey 2016*. In J. Bacher, A. Grausgruber, M. Haller, F. Höllinger, D. Prandner & R. Verwiebe (Hrsg.), *Sozialstruktur und Wertewandel in Österreich: Trends 1986–2016*. Springer, 515–531. https://doi.org/10.1007/978-3-658-21081-6_21 (Stand: 15.1.2024).
- Prandner, D. (2022). *Zu Datengrundlage und Datenqualität: Methodische Reflexion zur quantitativen Erhebung während der Corona-Krise*. In: Aschauer, W., Glatz, C., Prandner, D. (Hrsg.), *Die österreichische Gesellschaft während der Corona-Pandemie*. Springer, 325–347. https://doi.org/10.1007/978-3-658-34491-7_12 (Stand: 15.1.2024)
- Prandner, D. (2023). *Bevölkerung und Werte*. In K. Praprotnik und F. Perlot (Hrsg.), *Das politische System Österreichs*. Böhlau, 73–100.
- Prandner, D., Seymer, A. & Weichbold, M. (2023). *The new normal of online survey research: General population studies are going online*. In: R Sooryamoorthy und V Ojong Researching (Hrsg.), *the 'New Normal' Social World*. Routledge, 29–49.
- Schwartz, S. H. (2012). *An overview of the Schwartz theory of basic values*. Online Readings in Psychology and Culture, 2(1), 2307-0919.
- Vehovar, V. & Lozar Manfreda, K. (2017) *Overview: Online Surveys*. In Nigel G. Fielding et al. (Hrsg.) *The SAGE Handbook of Online Research Methods*. SAGE Publications, 143–161.